MULTIMEDIA UNIVERSITY®

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2017/2018

HPB2021 – BIOINFORMATICS PROGRAMMING II

(All sections / Groups)

26 OCTOBER 2017
9:00-11:00 AM
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This question paper consists of 4 pages including cover page.
2. Attempt all **FOUR** questions. All questions carry equal marks (10), and the distribution of the marks for each question is given.
3. Please write all your answers in the Answer Booklet provided.

QUESTION 1 [10 marks]

a) Write down the expected output of the following codes.

Python

i. [1 mark]

```
def fun(x, y):
    print x + y,
    print x * y

fun("2", 3)
fun("2", "3")
```

ii. [0.5 mark]

```
def what(n):
    if n == 0:
        return 0
    else:
        return n + what(n-1)
print what(5)
```

iii. [1 mark]

```
s = "Alabama and Alaska"
s[ : 7]
s[ 1 : 7 : 2]
s[ 2 : 7 :-1]
s[ : : -1]
```

iv. [1 mark]

```
text = raw input("Input text: ")
vowels = ['a', 'e', 'i', 'o', 'u', 'y' ]
for c in text.lower():
    if not c.isalpha():
        continue
    if c in vowels:
        print c,
        continue
    print "%sO%s"%(c, c),
```

Input text: The light was yellow, sir.

v. [1 mark]

```
m = [ [0] * 3 ] * 3
print m
m[0][2] = 7
print m
```

vi. [0.5 mark]

```
x= "Hello "
y= "World! "
x, y = y, x
print x+y
```

vii. [1 mark]

```
fred = [ 'Alice', 'goes', 'to', 'market' ]
fred.extend([ 'with', 'Mike' ])
last = fred.pop()
fred.append('Fred')
print 'A:', fred
fred = [ 'On', 'Tuesday,' ] + fred
print 'B:', fred
fred.reverse()
print 'C:', fred
fred.sort()
print 'D:', fred
```

viii. [0.5 mark]

```
for y in range(25, 20, -1):
    if y != 25:
        print y,
```

Continued .....

R Programming	
i. > x <- c(1:10) [1 mark] > x[(x>8) | (x<5)] > x > 8 | x < 5	ii. > smoker=matrix(c(70,50,65,140), ncol=2,byro=FALSE) > rownames(smoker)=c("male","female") > colnames(smoker)=c("smoke","nosmoke") > smoker= as.table(smokers) > smokers [1mark]
iii.> seq(3,30,length=10) [1 mark]	iv. > x=1:6 [0.5 mark] > y=cut(x,3) >y

QUESTION 2 [10 marks]

a) Given a string S as below

>>> S = "AAL532906 aaaatagtcaaatatatcccaattcagtatgcgctgagta"

Write Python code to search for r1 = "AGT" (uppercase) in S. If it is present, print "found", otherwise print "not found"
Write Python code to find and print the starting index (location) of r1 in S. If there is more than one, print all of them.
Write Python code to get and print the substring of S between 2 occurrences of r1.

[5 marks]

b) Write an R program to read a sequence from the user. The program counts the number of bases, "A","T","C", and "G" prints the length of the sequence, bases, and the number of occurrences. [4 marks]

c) Write R code to draw a histogram of table D$wg with graph title "Weight gain", x and y axis labels "Gain", and "Freq" and red color bars. [1 mark]

QUESTION 3 [10 marks]

a) Read all the sequences in a fasta file "example.fasta" using BioPython. Display the IDs and the sequences. Store the IDs in idList[], and the sequences in seqList[]. [5 marks]

b) Write BioPython code to create a Seq object of DNA using the IUPAC.unambiguous_dna alphabet find the number of 'G' and 'C' nucleotides in the sequence calculate the percentage of 'G' or 'C' in the sequence. [5 marks]

Continued.....

QUESTION 4 [10 marks]

a) Write a Python program to ask a user to input a DNA sequence. The program searches for restriction sites (in uppercase letters) in the input sequence. If a restriction site is present in the sequence, it prints the site name and the index of the starting point. If it is not present, then it prints the site name and "not present".

[5 marks]

Hint : here are sample restriction sites

```
"GAATTC",   # EcoRI
"GGATCC",   # BamHI
"AAGCTT",   # HindIII
```

b) What is a Dictionary in Python? Give an example for the following.

Create a Dictionary named *myL1* with at least 3 items
Add an item to the Dictionary
Remove an item from the Dictionary
List the keys in the Dictionary
Create a Dictionary *myL2* which contains the *myL1* Dictionary

[5 marks]

END OF PAPER